

The role of morphology in ordering verb-adjacent clitics: from syntax to prosody in Bulgarian and Tagalog

Loren Billings and Abigail Konopasky

This paper explores the hypothesis that the largely prosodic ordering of clitics in Tagalog is an innovation which has emerged from a stage in which the clitics were ordered by syntactic mechanisms. Now nearly no syntax is involved in the relative order of clitics. This Austronesian language and its close relatives appear to be unique in comparison to other documented clitic systems in this regard: namely, the Romance and Slavic subgroups of Indo-European. For this reason, closely related languages about which there is sufficient descriptive work have been investigated in order to determine if intermediate stages are attested. Preliminary findings presented below suggest that at least one such language, Cebuano, exhibits a mixed system in which both syntax and prosody are involved. The other language which we investigated, Bikol, shows properties between those of the other two languages. Such data support our hypothesis.¹

This paper is organized along ontogenic lines. We start by summarizing the claims in our previous work, which compares Tagalog with a very similar system in the Slavic language Bulgarian. We then present the comparative data among several Central Philippine languages which are not

¹ This work was part of a grant from the National Science Council, Republic of China (NSC 90-2416-H-126-015). Thanks also to several colleagues for their assistance: Edith Aldridge, Christine Alito, Leonard Babby, Kjetil Rå Hauge, Daniel Kaufman, Iliyana Krapova, Paul Kroeger, Mariana Lambova, Celeste Lee, Raphael Mercado, Roumyana Pancheva, Charles Randriamasimanana, Norvin Richards, Ian Roberts, Anna Roussou, Catherine Rudin, Ivanka Schick, Anne Sturgeon, Peter Svenonius, Keri Taylor, and Adam Werle. However, only we are to blame for any errors herein.

inconsistent with our hypothesis: Tagalog diachronically arose from a stage with syntactic arrangement of its clitics.

1 Background

Our collaboration has consisted of four papers, each to appear in a proceedings volume. In part because this is the last of the four to be submitted for publication, we do not wish to repeat all of our data or arguments here.² For this reason, this introductory section merely recaps our general arguments. We exemplify only those aspects relevant to the discussion in section 2 below. Our background discussion begins by demonstrating the significance of verb-adjacent clitics in revealing how clitics become affixes diachronically. We then sketch our theoretical model of the syntax-morphology interface. Finally, we show how the prosody can be called upon—namely, in Tagalog—to order elements which the syntax and morphology are unable to arrange temporally.

1.1 Verb-adjacent clitics: a window into the morphology

Much recent work in linguistics (e.g., Ura 2000; Nunes 2001) has dealt with the extent to which syntax manipulates morphological features. Clitics have been central to this work because they tend to progress historically from independent words into clitics and later affixes. The most closely examined languages in this regard have been from Ibero-Romance; see, for example, Nishida (1996) and Wanner (1996) as well as the references therein for further details.

² The remaining three papers discuss other, more synchronic aspects of the Bulgarian-Tagalog comparison: Billings & Konopasky (to appear-b) specifies and supports our theoretical claims about the conditions under which the syntax can and cannot require a particular order of clitics in the morphology. Billings & Konopasky (to appear-c) then sketches an explicit model of how the prosody takes over in Tagalog if the morphosyntax fails to require an order. In addition, in Billings & Konopasky (to appear-a) we discuss some consequences of our model relevant primarily to Bulgarian.

- (1) a. El sabe [que *lo* yo deseo]. *Old Spanish*
 he knows that it I desire
 ‘He knows that I desire it.’ [Nishida 1996: 335]
- b. (... que) yo también *lo* quiero. *Modern Spanish*
 ‘(... that) I also want it.’ [cf. *(...) yo *lo* también quiero]

(In the numbered examples clitics are shown in italics and most verbs are underlined.) Several of the Ngumpin languages, a subgroup of the Pama-Nyungan family spoken in Australia, have also been investigated (McConvell 1996).

A primary insight of this work is that clitic clusters show non-syntactic idiosyncrasies such as the Person-Case Constraint (which disallows third-person indirect objects from coexisting with first- or second-person direct objects in the cluster, not exemplified here) and consecutive homophony (Bonet 1991, 1995; Franks & King 2000). In (2a–b) the clitics appear in isolation, but when they co-occur, the expected surface form **le lo* in (2c) is not attested. Instead we find the form shown in (2d), with *se lo* instead.

- (2) a. A Pedro, *le* dieron el premio ayer. *Spanish*
 to Pedro 3.Dat gave.3Pl the pri[z]e yesterday
- b. El premio, *lo* dieron a Pedro ayer.
 the pri[z]e 3.Acc gave.3Pl to Pedro yesterday
- c. *A Pedro, el premio, *le lo* dieron ayer.
- d. A Pedro, el premio, *se lo* dieron ayer.
 ‘They gave the pri[z]e to Pedro yesterday.’
 [Bonet 1995: 608]

In (2a–b) the clitics *le* and *lo* double the preceding indirect- or direct-object nominals, respectively. However, nominals appearing after the verb in (2a–b) are not doubled by clitics.

Linguists (e.g., Terzi 1999) have tried to explain these and similar idiosyncrasies using strictly syntactic approaches. However, the mechanisms used become bulky; at some point in the diachronic shift from clitic- to affix-hood the syntax must relinquish control over the cluster's internal order. Because they are undergoing this diachronic progression, languages with verb-adjacent clitics offer a look inside this historical process. We therefore investigate Bulgarian (South Slavic, Indo-European) and Tagalog (Central Philippine, Austronesian) in particular because, while the Bulgarian cluster shows strictly syntactic ordering of its clitics, the cluster in Tagalog shows primarily prosodic sequencing.

Examples from Bulgarian and Tagalog illustrating this contrast in the ordering of clitics are shown in (3) and (4):

- (3) a. Az *ti* *gi* dadox. *Bulgarian*
 1Sg.Nom 2Sg.IO 3Pl.DO gave.1Sg
 'It's me that gave them to you.'
- b. Dadox *ti gi*.
 'I gave them to you.'
- (4) a. Kahapon *ka* *ba nila* nakita? *Tagalog*
 yesterday 2Sg.Sbj Q 3Pl.DO be.seen
 'Was it yesterday that they saw you?'
- b. Nakita *ka ba nila* (kahapon)?
 'Did they see you (yesterday)?'

In (3a) and (4a) the verb follows the cluster only if another word precedes it, whereas in the corresponding (b) examples the clitic cluster follows the initial verb.³ In both languages, the clitics and verb must be adjacent. However, the Bulgarian

³ The cluster keeps the same order in both languages regardless of whether the verb precedes or follows it. Greek, with free post-verbal order (Terzi 1999), is different.

clitic cluster shows syntactic ordering (i.e., indirect before direct object), while the Tagalog cluster shows primarily prosodic ordering (i.e., a monosyllabic pronoun before a disyllabic one, with a discourse particle in between).

1.2 The syntax-morphology interface

In this subsection we sketch our model of how the syntax feeds the morphology and ordinarily entails a specific order of the verb and the adjacent clitics. However, some syntactic configurations prohibit such ordering: namely, in Tagalog.

We assume a conventional T-model of generative grammar, with one leg consisting of Spell-Out followed by the morphological component (MC), then the phonological/prosodic components, and terminating in phonetic form (PF). We rely on two fundamental notions of grammar to account for all the data in Bulgarian: the EPP- or D-feature (as interpreted specifically by Alexiadou & Anagnostopoulou 1998) as well as the Linear Correspondence Axiom (LCA)—which states that if any element *A* c-commands some element *B* asymmetrically, then *A* precedes *B* in PF (Kayne 1994). In clauses like (3a) above, the subject moves over the clitics to SpecTP to check the D-feature in T. However, if no subject or topic is attracted to SpecTP, as in (3b), the verb itself checks this feature in T, raising over the clitics. We also adopt a strictly local approach to Economy (Collins 1997): a syntactic element may skip over a like element as long as it possesses the closest matching feature. As a result, when the syntax spells out to the PF leg, each element—the clitics, the verb, and any topic in SpecTP—is in its own projection and thus in an asymmetric c-command relation. This allows the syntax to dictate the linear ordering to the MC.

In Tagalog, on the other hand, pronominal clitics are base-generated as arguments within *v*P. Unlike in Bulgarian,

they are real pronouns and move to their respective agreement heads in order to check features.⁴ Consider (5):

- (5) Ipinakilala *ka* *nila* ... *Tagalog*
 be.introduced 2Sg.Sbj 3Pl.DO
 ‘They introduced you ...’

Our proposed syntactic structure, prior to inserting TP, is illustrated in (6a). To simplify the structures, we do not show traces; (6a) is the result of movement into the Agr heads. In addition, “|” indicates no precedence relation between items.

- (6) a. Syntax prior to TP insertion *Tagalog*
 [{2Sg|AgrS} [{3Pl|AgrDO} [ν P ... V ...]]]
 b. V-to-T movement in order to satisfy the D-feature
 [{ {V| {3Pl|AgrDO}} | {2Sg|AgrS} } | T } [AgrSP ...]]
 c. Result of the LCA and lexical insertion
 { ipinakilala | *nila* | *ka* } ...
 [+ V] [+ af_V] [+ af_V]

Because of multiple adjunction to either of the Agr heads in (6b), no asymmetric c-command relation obtains once the verb moves through. This configuration therefore prohibits the MC from ordering any of the elements in (6b). We emphasize, however, that the only way that the MC can fail to order elements is *within* the same morphological word; the LCA still always orders two matrix morphological words; the example in (5) and (6) show only one morphological word.

⁴ We offer two reasons to propose movement from argument positions. First, indirect-object pronouns surface in a PP, outside the clitic cluster (not shown here); a non-clitic word can precede such a PP. This suggests that the subject and direct-object pronouns are likewise base-generated in ν P. In addition, Tagalog clitics do not double overt nominals, as in Bulgarian and above in (2). Because of prior adjunction to Agr heads, the LCA thus cannot determine linear order because V later adjoins to the same heads.

Thus, the same mechanisms account for clitic-ordering differences in Bulgarian and Tagalog, despite the languages' differences. If the syntax results in structures with asymmetric c-command, as in Bulgarian, then the MC can order the elements. However, if there is no clear hierarchical relationship in the syntax, as in Tagalog, then the syntax and morphology fail to predict the order of the verb and clitics.

1.3 Prosodic ordering if the LCA cannot arrange the clitics

If there is even one clitic pronoun in a Tagalog clause, because of multiple adjunction to the agreement head, then the order of the verb and clitic(s) cannot be resolved by the syntax or morphology. As a result, prosodic factors take over and decide the surface order. In this subsection we begin with the empirical facts of how the clitics and verb are ordered relative to each other and the verb. We then briefly sketch our previous proposal, from Billings & Konopasky (to appear-c), which attempts to explain this ordering.⁵

Tagalog's clitic cluster is sensitive to three types of clitics: one-syllable pronouns, discourse particles, and two-syllable pronouns. They must appear in that order. We begin with the ordering of pronominal clitics, where syllabic weight determines the order. Next we show how discourse particles appear between the two types of clitic pronouns. We then discuss what happens if more than one of any of these three types of clitic appears in the same clitic cluster.

⁵ We are unable to go into the position of the clitic cluster relative to the rest of the clause in detail here. In both Bulgarian and Tagalog, the cluster precedes the verb unless there is no material in front of the verb, in which case the verb precedes the clitic(s). This was also true of medieval periods of Romance languages, discussed by Wanner (1996). During the past hundred-odd years, Bulgarian clitics have become unambiguously verb-adjacent, no longer requiring a full orthotonic word in front of them; an unstressed conjunction will do. This indicates that such clitics are no longer peninitial. In Tagalog, however, the clitics are apparently required to be both verb-adjacent and peninitial in their intonation phrase. As mentioned in a preceding footnote, the cluster's internal order is the same both pre- and post-verbally.

First, there can be up to two clitic pronouns in a clause, corresponding to the AgrS and AgrDO heads; cf. (6a–b). By contrast, in Tagalog AgrIO pronouns are in prepositional phrases and are not positioned as clitics. The full paradigms of subject- and object-pronominal clitics are presented in (7):

(7)	Clitic pronouns ⁶				<i>Tagalog</i>
		[± plural]	[± me]	[± you]	Subject Object
a.	1Sg	–	+	–	akó ko
b.	2Sg	–	–	+	ka mo
c.	3Sg	–	–	–	siyá niyá
d.	1Pl Incl.	+	+	+	táyo nátin
e.	1Pl Excl.	+	+	–	kamí námin
f.	2Pl	+	–	+	kayó ninyó
g.	3Pl	+	–	–	silá nilá

(One-syllable clitics are highlighted for clarity.) Note that no combination of syntactic features can describe them to the exclusion of the disyllabic clitics. If any of these three co-occurs with a disyllabic clitic pronoun, then the monosyllabic clitic must be initial in the cluster, as shown in (8a–c):

- (8) a. Nakita **ko** **siya**. *Tagalog*
 be.seen 1Sg.DO 3Sg.Sbj
 ‘I saw him/her.’ [Schachter & Otnes 1972: 185]
- b. Nakita **mo** **ako**.
 be.seen 2Sg.DO 1Sg.Sbj
 ‘You (Sg.) saw me.’ [Schachter & Otnes 1972: 185]
- c. Nakita **ka** **nila**.
 be.seen 2Sg.Sbj 3Pl.DO
 ‘They saw you (Sg.).’

⁶ Schachter (1973: 88) lists a pair of archaic disyllabic clitic pronouns which render dual number; their inclusion would not affect our claims below. Object clitic pronouns can also be possessive; such forms co-occur with subject clitics in copular clauses.

No orders other than these are acceptable; see also (4a–b) and (5) above. Thus, clusters of pronominal clitics with heterogeneous syllabic weight are rigidly ordered. (We put off homogeneous pronoun-clitic clusters for the time being.)

The third main type of clitic with regard to positioning are the so-called discourse particles. Examples of these, corresponding to (8a–c), are shown in (9a–c), respectively:

- (9) a. Nakita *ko* ***na*** *siya.* *Tagalog*
 be.seen 1Sg.DO already 3Sg.Sbj
 ‘I saw him/her already.’
- b. Nakita *mo* ***yata*** *ako.*
 be.seen 2Sg.DO perhaps 1Sg.Sbj
 ‘Perhaps you (Sg.) saw me.’
- c. Nakita *ka* ***ba*** *nila?*
 be.seen 2Sg.Sbj Q 3Pl.DO
 ‘Did they see you (Sg.)?’

Regardless of its syllabic weight, in Tagalog any particle must appear after any one-syllable clitic pronoun and before any two-syllable clitic pronoun in the same clitic cluster.

The next issue is what happens in clusters with clitics of the same type. If more than one of any type of clitic is present, then quite different rules apply. We consider all three possibilities, beginning at the end of the cluster.

If two disyllabic pronominal clitics co-occur in a cluster, then no categorical restrictions apply. In passive clauses at least, there is merely a preference for the object clitic to precede the subject clitic. For example, Schachter & Otnes (1972: 185) describe the order in (10b) as “highly unusual”; Schachter (1973: 218) adds that usually Obj.3Sg *niya* precedes any disyllabic subject pronoun: “This is, however, only a matter of preference, and a sequence such as *ako niya*

is grammatical, although *niya ako* is preferred.” We have added the superscripted question mark in front of (10b).

- (10) a. Nakita *niya* *ako*. *Tagalog*
be.seen 3Sg.DO 1Sg.Sbj
- b. [?]Nakita *ako niya*.
‘He/She saw me.’ [Schachter & Otanes 1972: 185]

This optionality stands in stark contrast with the inviolable restrictions on heterogeneous clusters, in (8) and (9) above.

Our account relies on syntactic and pragmatic properties, although crucially indirectly by way of accentual properties: Kroeger (1993: 119) observes that pronominal clitics with the role of Actor tend to appear first. (In Tagalog and related languages, the label Actor is used to mean something like ‘protagonist’; causation or intent usually associated with the more familiar role of Agent are not necessarily present.) Although Kroeger does not make the connection, his observation appears to be epiphenomenal to one of the rules of thumb elsewhere in his study about the ordering of *non-clitic* (post-verbal) arguments—expanding on Schachter & Otanes (1972: 83–85) somewhat: “the Actor phrase tends to precede all other arguments”; another tendency, he reports, is for the subject argument to follow all other nominals (Kroeger 1993: 111). In a passive clause both of these tendencies are satisfied by object-subject order, as in (10a). However, regarding clauses with active voice, the order of non-clitic nominal expressions is not generally agreed upon in the literature on Tagalog word order (Kroeger 1993: 112).

Nor, it appears, do disyllabic clitic pronouns show any particular ordering tendency in clauses with active-voice verbs. We are unaware of any published mention of this

point.⁷ We assume without further discussion that the ordering of both disyllabic clitic pronouns and full nominals is related. There appear to be independent Actor-first and subject-last effects going on in the language.

What formal sense is to be made of the two tendencies? Clearly, subject-last is at least in part syntactic, since the external argument is generally associated with distinct arboreal positioning.⁸ It remains a mystery, however, why the most arboreally superior argument is final rather than initial. The other tendency, Actor-first, is likely a misnomer; semantic (= thematic) roles should probably not be identified in the syntax proper (Grimshaw 1991). Instead, it is this argument's pragmatic prominence, in terms of its point of view, that is the likely source of the Actor-first effect.

Next, a conceptual problem arises because both clitic and non-clitic nominals are affected by the same ordering tendencies. If there is no phonological information in the LF leg of the T-model, then there should be no reference to meaning (including pragmatic notions) in the PF leg as well. Indeed, we argue above (in §1.2) that the syntax holds no sway over cluster ordering in Tagalog. How then can pragmatic and syntactic information be relevant to ordering the cluster? Our solution is for the aforementioned syntactic and pragmatic properties of subjects and Actors to be mapped first into prosodic properties, which in turn entail specific orders when it comes time in the derivation to determine the order of two disyllabic clitic pronouns. Namely, the subject is emphasized somehow in the syntax and this prominence is transferred during Spell-Out. Similarly, the Actor must be

⁷ Such discussion in the literature on Tagalog clitics uses passive clauses exclusively.

⁸ The fact that only subjects must carry a definite interpretation in Tagalog (Schachter 1976), and that definiteness further entails presupposed or non-focal pragmatics, is irrelevant here. Being equally definite, all pronouns are equal in presuppositionality.

de-emphasized in both ways.⁹ The presence or lack of such *prosodic* prominence is the crucial factor affecting the order of two disyllabic clitic pronouns in Tagalog at PF. Below in this subsection we show how such factors affect ordering.¹⁰

Moving next to how two or more discourse particles are ordered, Schachter & Otones (1972: 413–415) and Schachter (1973: 219–220) report a complex matrix of subclasses in which syllabic weight tends to be relevant, with monosyllabic members generally preceding disyllabic ones. This matrix is mainly tendencies, however, without the rigidity shown in (8) or (9) above. Our model (discussed below) relies on post-prosodic insertion of particles, after the pronominal clitics' grouping into prosodic words has been determined. As such, particles could well be ordered using a templatic approach.

This brings us to what occurs when two monosyllabic pronouns co-occur in the clitic cluster. Although there are three such clitics (namely: Obj.1Sg *ko*, Obj.2Sg *mo*, and Sbj.1Sg *ka*), no two clitics of this type can co-occur in the same clause for various reasons. First, *ko* and *mo* are both direct objects and thus in complementary distribution. Similarly, *mo* and *ka* are both 2Sg, requiring a non-clitic reflexive pronoun in place of one of the two clitics. The only other combination among them is *ko* and *ka*. However, no combination of these two clitics is possible. Instead, similarly to the situation in Romance illustrated above in (2), a suppletive portmanteau clitic *kita* is attested instead, in (11c):

⁹ Similarly, Billings & Rudin (1996) propose an interaction of Superiority- and animacy-based constraints to determine the order of multiply fronted *wh* phrases in Bulgarian. The constraints used there would appear to apply only at LF. However, whichever mechanism requires all the *wh* phrases to front in overt syntax would probably also entail the same ordering of all interrogative phrases at PF as well.

¹⁰ How the non-clitic nominals are ordered based on such prominence remains open. Some literature (summarized by Bošković 2001: 94, 255) argues for movement at PF of even non-clitic material (e.g., heavy-NP shift in English). For the present paper's purposes, we assume that syntactic movement is preferable for non-clitic nominals.

- (11) a. *Nakita *ko ka*. Tagalog
 be.seen 1Sg.DO 2Sg.Sbj
- b. *Nakita *ka ko*.
- c. Nakita *kita*.
 be.seen {1Sg.DO|2Sg.Sbj}
 ‘I saw you (Sg.)’ [Schachter & Otanes 1972: 185]

This portmanteau clitic appears in the same order as any other disyllabic clitic in the cluster; particle clitics precede disyllabic pronominal clitics, as in (9a–c) above. Namely, *kita* invariably follows such particle clitics: *Nakita ba kita?* ‘Did I see you (Sg.)?’ This ordering shows that *kita* is a single clitic, ordered after *ba* based solely on its syllabic weight. Schachter (1973: 221–223) uses such data to argue against transformational-syntactic approaches to the ordering of clitics; his argument remains valid with regard to more recent approaches which are limited to exclusively syntactic mechanisms to account for the ordering of clitics.

The special positioning of *kita* is therefore formidable evidence against syntactically driven ordering of clitics. By contrast, Bulgarian exhibits no such suppletion in its clitic clusters. This suggests to us that syntax orders the clitic cluster in Bulgarian, whereas other factors dictate the linearization of the clitics and verb in Tagalog.

The main constituent order that needs to be explained is shown in (9a–c) above. In our view, there are two theoretical challenges: the ordering of the pronominal clitics according to their syllabic weight and then placing the discourse particles between one- and two-syllable clitics. We sketch our approach to these issues here briefly; see Billings & Konopasky (to appear-c). First, lacking any ordering in the output of the MC, Optimality-theoretic constraints serve to

select the best-formed arrangement of the verb and adjacent clitic *pronouns* in a given clause. One constraint restricts pronominal clitics from being initial by requiring them to follow an accented word. Another constraint requires them to be as early in the clause as possible. Yet another constraint requires disyllabic clitics to form their own prosodic word. The three together result in the cluster preceding the verb if there is a preceding orthotonic verb; otherwise, the cluster follows the verb. Moreover, monosyllabic pronouns precede disyllabic ones.¹¹ In addition, the position of particle clitics crucially depends on how the other two types of clitic pronouns have been organized into prosodic words.¹² As such, particles are inserted far later, post-prosodically. This results in the rather peculiar sandwiching of particles between heterogeneous pronominal clitics, in (4) and (9a) above.

To conclude this subsection, we have laid out the relevant ordering facts in Tagalog: with up to one monosyllabic clitic pronoun, followed by up to several discourse particles, then up to two disyllabic pronominal clitics. Various additional restrictions affect the internal ordering of each of these three clitic types. We also sketched briefly how such unusual, mostly prosodic ordering can be derived in our framework.

To summarize section 1, we have outlined our overall model of how verb-adjacent clitics are at a diachronic turning point. They are about to lose their syntactic status. For this reason, we investigate them in order to elucidate the interface of syntax and morphology. We have done so by showing that in Tagalog certain syntactic structures can result in

¹¹ In light of our discussion above about the order in which two disyllabic clitic pronouns appear, a fourth constraint requiring that prosodically prominent pronouns appear at the end of a phonological phrase would achieve the order in (10a). We have not addressed this particular ordering issue in Billings & Konopasky (to appear-c).

¹² This is strikingly similar to the positioning of *li*, the yes/no marker in Bulgarian. The prosodic-word groupings are determined before *li* is inserted (Rudin et al. 1999).

morphologically unorderable strings. In such cases prosodic constraints emerge to determine the surface ordering.

2 Comparison of Tagalog with Bikol and Cebuano

We turn now to the main findings of this paper: evidence from two closely related languages' clitic-ordering systems. One of these, Bikol, reveals some differences from Tagalog's system in its inventory of clitic pronouns. However, given these lexical differences, the clitic clusters are ordered almost identically in the two languages. The other language, Cebuano, shows major differences in both its inventory of clitics and their positioning. Following an overview of how these three languages are related to each other, we present the lexical facts by comparing the languages' inventories of clitic pronouns. We then move on to the ordering of these clitics within the cluster. The section ends by comparing various kinds of suppletive phenomena which speak to the essentially morphological status of the cluster in all three languages.

Studies agree—e.g., Zorc (1975)—that the smallest subgroup including Tagalog, spoken in Central and Southern Luzon, is Central Philippine (CP). The other members of CP for which good linguistic documentation is available include Bikol (spoken around the peninsula of the same name in Southeastern Luzon), Bisayan (over a dozen languages spoken mainly in the Visayas: smaller islands between Luzon and Mindanao), and Mansakan (around Southern Mindanao).

For this paper we consulted grammars of Bikol, specifically the variety in Naga City (described by Mintz 1971; 1973), and the variety of Cebuano spoken in Cebu City (Wolff 1965) as the representative of the Bisayan languages. We also refer to Zorc (1975), a comparative-historical investigation of Bisayan with a chapter on CP as well.

2.1 Comparing the inventories of clitics

Expanding the Tagalog table above in (7), we add the corresponding forms in Bikol (Mintz 1971: 12–13) and Cebuano (Wolff 1965: 64; Zorc 1975: 91–94), along with the reconstructed CP forms (Zorc 1975: 279) in the right margin:

(12) Subject clitics: *Tagalog* *Bikol* *Cebuano* *Proto-CP*

a. 1Sg	<i>akó</i>	<i>akó</i>	<i>qakú</i>	<i>ku</i>	* <i>akú</i>
b. 2Sg	<i>ka</i>	<i>ka</i>	qikáw	<i>ka</i>	* ikáw
c. 3Sg	<i>siyá</i>	<i>siyá</i>	<i>siyá</i>	<i>sya</i>	* <i>s-iyá</i>
d. 1Pl Incl.	<i>táyo</i>	kitá	kitá	<i>ta</i>	* kitá
e. 1Pl Excl.	<i>kamí</i>	<i>kamí</i>	<i>kamí</i>	<i>mi</i>	* <i>kamí</i>
f. 2Pl	<i>kayó</i>	<i>kamó</i>	<i>kamú</i>	<i>mu</i>	* <i>kamú</i>
g. 3Pl	<i>silá</i>	<i>sindá</i>	<i>silá</i>	—	* <i>s-idá</i>

(13) Object clitics: *Tagalog* *Bikol* *Cebuano* *Proto-CP*

a. 1Sg	<i>ko</i>	<i>ko</i>	nákuq	<i>ku</i>	* <i>-ku</i>
b. 2Sg	<i>mo</i>	<i>mo</i>	nímu	<i>mu</i>	* <i>-mu</i>
c. 3Sg	<i>niyá</i>	<i>niyá</i>	<i>níya</i>	—	* <i>-ña</i>
d. 1Pl Incl.	<i>nátin</i>	<i>ta</i>	nátuq	<i>ta</i>	* <i>-ta</i>
e. 1Pl Excl.	<i>námin</i>	<i>mi</i>	námuq	—	* <i>-mi</i>
f. 2Pl	<i>ninyó</i>	<i>nindó</i>	<i>nínyu</i>	—	* <i>-yu</i>
g. 3Pl	<i>nilá</i>	<i>nindá</i>	<i>níla</i>	—	* <i>-da</i>

As in (7), monosyllabic clitics are shown in italics.¹³ For expository purposes alone, we show in bold-face type certain salient differences from the corresponding Tagalog forms.

As the starred (i.e., reconstructed) forms also show, in some parts of the paradigm Tagalog is the exception. Most notably, in (12d) *kitá* is the Sbj-1PlIncl clitic pronoun in the remaining languages, not a portmanteau clitic as in (11c)

¹³ Mintz (1971) and Zorc (1975) show the stress for the disyllabic forms in Bikol and Cebuano. Wolff (1965) uses a mixture of stress and vowel length; we show Zorc's notation here and in the examples, from Wolff, in (25) and (26) below. Wolff and Zorc also use *q* to render the glottal stop. Lastly, in (12c) *sya* is from Zorc (1975: 94) only.

above. (See §2.3 for more on suppletive forms.) In addition, as (13d–e) show, Tagalog innovated both 1Pl object forms.

Of greater significance to this study is the comparatively richer inventory of one-syllable forms in the other two modern languages. As in Tagalog, Bikol’s inventory of monosyllabic pronouns primarily contains object clitics. Bikol attests monosyllabic object-1PlIncl and -1PlExcl forms. Cebuano attests even more monosyllabic forms with homophonous subject and object forms in both the 1Sg and 1PlIncl, as well as monosyllabic subject forms in (12c, e–f).

Moving next to disyllabic forms, Tagalog and (to a lesser extent) Bikol have innovated several two-syllable object clitics. Cebuano, on the other hand, attests a double paradigm, with a full paradigm of disyllabic pronouns.¹⁴

However, in all three modern languages the two-syllable forms are only optionally clitics. For example, the disyllabic subject forms in (12) can be clefted or topicalized, appearing clause-initially; such forms are not clitics. The crucial test for cliticness is as in (4a) above, following a fronted adjunct and before the verb. All three modern languages clearly attest the forms in (12) and (13) in such a position.¹⁵

To summarize briefly, Tagalog differs to varying degrees with Bikol and Cebuano in its inventory of clitic pronouns. More differences arise in the arrangement and combination (including suppletion) of the clitics in the three languages. These are discussed in the following two subsections.

¹⁴ In addition to the only one-syllable subject clitic, *ka*, Tagalog and Bikol retain non-clitic *ikáw/iká*, resp. (Schachter & Otones 1972: 88; Mintz 1971: 12–14), reflexes of Proto-CP *ikáw. In both languages, *iká(w)* is exclusively non-clitic in distribution. In Cebuano disyllabic forms are used as clitics even if a one-syllable counterpart exists. Wolff (1965: 219) implies that disyllabic pronouns may stand in for short forms to avoid prohibited sequences of one-syllable pronouns; see also §2.3 below.

¹⁵ Wolff (1965: 215), Schachter & Otones (1972: 493), and Mintz (1973: 167, 177, 227, 264, 277) list several such crucial examples of disyllabic subject clitics.

2.2 Comparing the arrangement of clitics

Bikol positions its clitics almost identically to Tagalog, whereas Cebuano shows a quite distinct arrangement. As with its inventory, Cebuano's ordering is considerably more complex, revealing more apparently syntactic effects.

Beginning with Bikol, although we have not found explicit discussion of how clitics are ordered, clauses with at least two clitic pronouns or particles in Mintz (1973)—in all, just about 100 examples—provide a corpus from which several generalizations can be drawn. First, if there is a monosyllabic pronoun, it precedes all other clitics. Next, if there is a disyllabic pronoun, it follows all particles. Thus, Bikol's system is basically like that of Tagalog.¹⁶

On the other hand, although similar to Tagalog and Bikol, Cebuano shows particularly informative distinctions:

(24) Arrangement of clitics: *Cebuano*

- a. The monosyllabic object pronouns: *ku*, *mu*, or *ta*
- b. The monosyllabic subject-2Sg pronoun: *ka*
- c. A specific set of (mostly monosyllabic) particles
- d. The monosyllabic subject-2Sg pronoun: *ka*
- e. The other particles
- f. The monosyllabic pronouns aside from *ka*
- g. The disyllabic pronouns

[simplified somewhat from Wolff 1965: 217]

In Cebuano, as in the other two languages, *ka* can appear early in the cluster, in accord with line (24b) and exemplified

¹⁶ Bikol shares two more complex issues with Tagalog, beyond the scope of this paper: split clusters and optional clisis, discussed by Schachter & Otones (1972: 184), Mintz (1973: 168, 277), and Kroeger (1993: 121–123). However, one apparent particle, *man giraray* 'still, once again', has unexplained initial positioning in the cluster (Mintz 1973: 168, 170, 184). In addition, suppletive *taká* (cf. §2.3 below) seems to be ordered as two monosyllabic pronouns (i.e., appearing before particles).

standpoint, but there is still no formal rationale for this order. In the next subsection we discuss various restrictions against multiple monosyllabic pronouns. The evidence, while murky, suggests there is just a single landing site for the clitic heads.

2.3 *Suppletive phenomena in Bikol and Cebuano*

We return to the type of lexical substitution observed for Tagalog above in (11c), with a portmanteau form used in place of either combination of *ko* and *ka*. Bikol and Cebuano have even more extensive data of this sort, likewise avoiding combinations of monosyllabic pronouns. These data suggest the following. First, morphology is involved in the clitic cluster's composition, supporting our assumptions about late insertion of PF features. In addition, a previous diachronic stage of placing clitics in peninitial position, motivated perhaps by checking a syntactic feature, might account for there being only one pronoun slot in the clitic cluster. These data thus provide a hint about the syntax of how clitics move syntactically in the modern languages. We begin, as above, with Bikol and move to the more complex Cebuano facts.

Bikol also resists having multiple monosyllabic clitic pronouns, but in more ways. As in (11c) above, if both Sbj-2Sg and Obj-1Sg features are present, Bikol uses the portmanteau form *taká*, as Mintz (1971: 14–15; 1973: 96–97) reports.¹⁸ In addition, Bikol idiosyncratically employs the non-clitic *iká* as the Sbj-2Sg pronoun, instead of clitic *ka*, if 1Plexcl *mi* is used. This, according to Mintz (1971: 14), avoids confusing **ka mi* with the Sbj-1Plexcl clitic *kamí*. Such replacement, it seems to us, might also avoid generating

¹⁸ As for the position of *taka*, Mintz (1973) lists no crucial data. We conducted a web search for clusters of *taka* and the common particle clitics in Bikol. We found several examples of *taka* preceding either of *lang* 'only' and *na* 'already', but found no examples of *taka* following these particles. This suggests that it is ordered as any monosyllabic would be, unlike the ordering of Tagalog's *kita*: as any disyllabic clitic.

multiple monosyllabic clitic pronouns. Indeed, our survey of the data with clitic clusters in Mintz (1973) revealed not even one example of two monosyllabic clitic pronouns.¹⁹ Given the Tagalog-like arrangement of the clitic cluster (discussed in §2.2 above), Bikol appears to circumvent such possible combinations (two in all) just as relentlessly as in Tagalog.

Similarly to Tagalog *kitá* and Bikol *taká*, Cebuano uses any of *tikáw*, *tiká*, or *ta ká* instead of Obj-1Sg *ku* and Sbj-2Sg *ka* (Zorc 1975: 90), as the examples in (26a–c) show.

- (26) a. tagáqan ***tikáw*** g gása *Cebuano*
 will.be.given {1Sg.DO|2Sg.Sbj} a present
- b. tagáqan ***tiká*** g gása
- c. tagáqan ***ta ká*** g gása
- ‘I will give you a present.’ [Wolff 1965: 219]

In addition, Cebuano replaces Sbj-2Pl *mu* and Obj-1Sg *ku* with *ta mu*. Two of the suppletive forms written as separate words above—*ta ka*, in (26c), and *ta mu*—can also be separated by a particle clitic such as the particle *man* ‘but’ (Wolff 1965: 219), unlike Tagalog and Bikol.²⁰ However, these suppletive forms do not exhaust all possible combinations of monosyllabic pronominal clitics in Cebuano; cf. (12) and (13) above. Without saying so explicitly, Wolff (1965: 219) implies that disyllabic forms are used to avoid combinations of monosyllabic pronouns (similarly to *ka + mi* in Bikol becoming *mi iká*, discussed above in this subsection). Furthermore, we have not uncovered any

¹⁹ We assume that the only other (non-reflexive) pairing in Bikol is semantically illicit. Namely, 1PlIncl *ta* in (13d) and 2Sg *ka* in (12b) both share the feature [+you].

²⁰ Our sampling suggests that *tiká* or *tikáw* are positioned as a disyllabic clitic (as in Tagalog, not Bikol): after at least some type-(24c) particles (e.g., *lang/lámang* ‘only’).

combinations of Cebuano monosyllabic clitic pronouns in the same cluster, even separated by particles (e.g., in Shryock 1993). For now, we assume that such combinations are illicit.

It seems, then, that all three languages have an identical restriction against two monosyllabic clitic pronouns in the same cluster. Because there appears to be no problem on the interpretive side of the T-model (i.e., Logical Form), the restriction appears to be at some point between Spell-Out and PF. This restriction is, furthermore, clearly not phonological, because monosyllabic particles appear next to monosyllabic pronouns and each other in all three languages. Even in Cebuano, where the two one-syllable pronouns could be separated by one or more particles, such combinations are apparently impossible.²¹ We therefore conclude tentatively that the restriction must be morphological in nature.

The comparative data in the foregoing subsection are furthermore consistent with a previous peninitial-positioning stage (Wackernagel 1892) in these languages' histories. Perhaps the subject, because of its A-bar status, was excluded from checking the D-feature in SpecTP. The direct object was thus the closest available argument in the tree (stopping through SpecAgrDOP on the way). Ambiguously both a phrase and a head, object *pronouns* then more economically moved into T itself. Over time, these pronouns eroded phonologically (Zorc 1975: 279). At this point the verb began to satisfy the D-feature. The object pronoun's reduction became reinterpreted as syntactic and later as morphological dependency to the verb. Even later, AgrS began to attract subject pronouns. This scenario is mainly speculation but at least consistent with what is known about

²¹ Our account (sketched in §1.3 above) of how particles in Tagalog are inserted post-prosodically might explain why even non-adjacent monosyllabic pronouns are illicit in Cebuano. They would have to be adjacent prior to the insertion of any particles.

how words become clitics historically. This D-feature-checking scenario also accounts for why only one type of pronoun, direct objects, appeared in the cluster at first.

This story is also consistent with our synchronic analysis of Tagalog. Because there is multiple adjunction to the AgrS and AgrDO heads, the syntax and morphology are unable to order the pronouns relative to the verb; the MC merely assures that the clitics remain adjacent to the verb, as affixes. Because of this sequence of historical stages, the cluster emerged with prosody as the only determiner of internal ordering. (In Bulgarian and other languages, however, the pronouns first became reinterpreted as object-agreement heads, allowing the syntax to dictate the surface order of the morphological constituents.) Bikol and Cebuano could easily be analyzed along similar theoretical lines as Tagalog.

To summarize section 2, we have compared the clitic clusters of two languages closely related to Tagalog in order to make sense of how the unusual order of clitics in that language emerged diachronically. Bikol shows only slight differences in ordering but attests a slightly larger inventory of monosyllabic pronouns and suppletive work-arounds. Cebuano, however, exhibits a significantly distinct system in inventory, ordering, and suppletion. Significantly, all three languages allow only a single monosyllabic pronoun.

3 Directions for further research

This working paper has shown preliminary evidence that Tagalog's ordering of clitics constitutes an innovation over an earlier, more syntactic proto-system. In this final section we discuss three additional areas where advances are needed.

First, more historical work is needed to solidify the speculative claims made here. It would help, for instance, to

investigate other subgroups of CP (namely, Mansakan). Neighboring, less closely related languages to the north, such as Kapampangan, might reveal areal evidence as well. There have also been nearly 500 years of written evidence since first European contact; philology might therefore reveal more recent changes. This paper, in many ways, is only a start.

In addition, a more thorough analysis of modern Tagalog is needed. While much descriptive work has been published, going back to Bloomfield (1917) and even the work of Spanish missionaries before him, there is still little consensus about the syntax of Tagalog and related languages. This was the tenor of a recent meeting of the Austronesian Formal Linguistics Association. Much of the discussion there, while careful and explicit, was by necessity mostly pre-theoretical.

Finally, linguists working on more familiar languages need to be more receptive to data that do not necessarily fit within the models applicable to the more familiar languages. Both of us, trained as Slavists, are familiar with anecdotes of how exotic material was unwelcome in generativist circles, more concerned with problems in, at first, English and then in, say, Germanic or Romance. By now, the Slavic languages are firmly represented in the generative enterprise, especially in the realm of clitics (e.g., Franks & King 2000 or Bošković 2001). However, having taken up the even less familiar Austronesian languages discussed above, we sense that linguists have now merely pushed back the pale of acceptability, but not beyond the limits of just a few familiar languages. The change needed is not to move an imaginary boundary but to discard the attitude that some languages are still unworthy of serious investigation. We therefore hope that this work will urge as well as inform our colleagues.

References

- Alexiadou, A. & E. Anagnostopoulou (1998) "Parameterizing Agr: word order, V-movement, and EPP-checking." *Natural Language and Linguistic Theory* 16: 491–539.
- Billings, L. & A. Konopasky (to appear-a) "Further arguments for morphology: clitics and the shift to affixhood." *Annual Workshop on Formal Approaches to Slavic Linguistics* [11]. *The Amherst meeting*. Ann Arbor: Michigan Slavic Publications.
- Billings, L. & A. Konopasky (to appear-b) "Morphology's role in ordering verb-adjacent clitics." In L. Mikkelsen & C. Potts (eds.), *Proceedings of the 21st West Coast Conference on Formal Linguistics*, Somerville, MA: Cascadilla Press, 29-42.
- Billings, L. & A. Konopasky (to appear-c) "Reassessing the Role of Syntax inside the Morphological Word: Verb-Adjacent Clitics in Tagalog and Bulgarian." In A. Riehl & M.T. Savella (eds.), *Cornell Working Papers in Linguistics* 19 (= Proceedings of the 9th Austronesian Formal Linguistics Association).
- Billings, L. & C. Rudin (1996) "Optimality and Superiority: A new approach to overt multiple WH-ordering." In J. Toman (ed.), *Annual Workshop on Formal Approaches to Slavic Linguistics* [3]. *The College Park meeting*. Michigan Slavic materials 38. Ann Arbor: Michigan Slavic Publications, 35-60.
- Bloomfield, L. (1917) "Tagalog texts with grammatical analysis. [Part 2 Grammatical analysis.]" *University of Illinois Studies in Language and Literature* 3(3).
- Bonet, E. (1991) *Morphology after syntax: pronominal clitics in Romance*. Massachusetts Institute of Technology Ph.D. dissertation.
- Bonet, E. (1995) "Feature structure of Romance clitics." *Natural Language and Linguistic Theory* 13: 607–647.
- Bošković, Ž. (2001) *On the nature of the syntax-phonology interface: cliticization and related phenomena*. Amsterdam: Elsevier.
- Collins, C. (1997) *Local Economy*. Cambridge, MA: MIT Press.
- Franks, S. & T.H. King (2000) *A handbook of Slavic clitics*. New York: Oxford University Press.
- Grimshaw, J. (1991) *Argument structure*. Cambridge, MA: MIT Press.
- Kayne, R.S. (1994) *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Kroeger, P. (1993) *Phrase structure and grammatical relations in Tagalog*. Stanford, CA: CSLI Publications.
- McConvell, P. (1996) "The functions of split-Wackernagel clitic systems: pronominal clitics in the Ngumpin languages (Pama-Nyungan family, Northern Australia)." In A.L. Halpern & A.M. Zwicky (eds.), *Approaching second: second position clitics and related phenomena*. Stanford, CA: CSLI Publications, 299–331.
- Mintz, M.W. (1971) *Bikol grammar notes*. Honolulu: University of Hawaii Press.

- Mintz, M.W. (1973) *Case and semantic affixes of Bikol verbs*. Diss. University of Hawaii.
- Nishida, C. (1996) "Clitics in Old Spanish and categorical grammar." In A.L. Halpern & A.M. Zwicky (eds.), *Approaching second: second position clitics and related phenomena*. Stanford, CA: CSLI Publications, 333–373.
- Nunes, J. (2001) "Sideward movement." *Linguistic Inquiry* 32: 303–344.
- Rudin, C., C. Kramer, L. Billings, & M. Baerman (1999) "Macedonian and Bulgarian *li* questions: beyond syntax." *Natural Language and Linguistic Theory* 17: 541–586.
- Schachter, P. (1973) "Constraints on clitic order in Tagalog." In A.B. Gonzalez (ed.) *Parangal kay Cecilio Lopez*. Manila: Linguistic Society of the Philippines, 214–231.
- Schachter, P. (1976) "The subject in Philippine languages: topic, actor, actor-topic, or none of the above." In C.N. Lee (ed.) *Subject and topic*. New York: Academic Press, 491–518.
- Schachter, P. & F.T. Otnes (1972) *Tagalog reference grammar*. Berkeley: University of California Press.
- Shryock, A. (1993) "A metrical analysis of stress in Cebuano." *Lingua* 91: 103–148.
- Terzi, A. (1999) "Clitic combinations, their hosts, and their ordering." *Natural Language and Linguistic Theory* 17: 85–121.
- Ura, H. (2000). *Checking theory and grammatical functions in universal grammar*. New York: Oxford University Press.
- Wackernagel, J. (1892) "Über ein Gesetz der indogermanischen Wortstellung." *Indogermanische Forschungen* 1, 333–436.
- Wanner, D. (1996) "Second position clitics in medieval Romance." In A.L. Halpern & A.M. Zwicky (eds.), *Approaching second: second position clitics and related phenomena*. Stanford, CA: CSLI Publications, 537–578.
- Wolff, J.U. (1965) *Cebuano Visayan syntax*. Diss. Yale University.
- Zorc, R.D.P. (1975) *The Bisayan dialects of the Philippines: subgrouping and reconstruction*. Diss. Cornell University.

Loren Billings
 Department of Foreign Languages and Literature
 National Chi Nan University
 Puli, Nantou, Taiwan 545 Republic of China
 billings@ncnu.edu.tw

Abigail Konopasky
 English Department
 University of New Orleans, Lakefront
 New Orleans, LA 70148 U.S.A.
 akonopas@uno.edu